DOCUMENT RESUME

ED 315 763 CS 212 144

AUTHOR Brody, Pamela; And Others
TITLE The Current State of Assessment in Literature. Report Series 3.1.
INSTITUTION Center for the Learning and Teaching of Literature, Albany, NY.
SPONS AGENCY National Endowment for the Arts, Washington, D.C.; Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 89
GRANT G008720278
NOTE 57p.
AVAILABLE FROM Center for the Learning and Teaching of Literature, School of Education, 1400 Washington Ave., ED B-9, University of Albany, State University of New York, Albany, NY 12222 ($7.00).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Content Analysis; *Educational Assessment; *Evaluation Methods; *Literature; National Surveys; Reading Comprehension; Secondary Education; Student Evaluation; Test Format

ABSTRACT

A study examined state-mandated literature assessment procedures and undertook a content analysis of tests produced by proprietary testing companies and textbook producers in order to ascertain how literature achievement is assessed in the United States. Results indicated that 61% of the school systems assess literature as part of an assessment of reading and that the main focus of these assessment programs is on comprehension rather than cultural literacy. The content analysis indicated that a majority of items in proprietary testing programs deal with comprehension of the content of the text rather than any literary aspects such as language, structure, or voice and that tests in commercial anthology series also focus on comprehension. The results also indicated that the one exception to the overall assessment picture occurred in university placement examinations which contained a higher number of items calling for knowledge of literature and literary skills. (Six tables and 15 figures of data are included; 16 references, an interview schedule, a list of commercial achievement tests reviewed, a classification scheme for test items, and 8 tables of data are attached.) (Author/RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: CS

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. TM

ED315762

The Current State of Assessment in Literature

Pamela Brody
Carol DeMilo
Alan C. Purves

Center for the Learning & Teaching of Literature

University at Albany
State University of New York
School of Education
1400 Washington Avenue
Albany, New York 12222

CS212144

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The Current State of Assessment in Literature

Pamela Brody
Carol DeMilo
Alan C. Purves

Center for the Learning and Teaching of Literature
University at Albany
State University of New York
1400 Washington Avenue, Albany, New York 12222

Report Series 3.1

1989

Preparation of this report was supported in part by grant number G008720278, which is cosponsored by the U.S. Department of Education, Office of Educational Research and Improvement (OERI/ED), and by the National Endowment for the Arts (NEA). However, the opinions expressed herein do not necessarily reflect the position or policy of OERI/ED or NEA, and no official endorsement of either agency should be inferred.

Center for the Learning and Teaching of Literature

The Center for the Learning and Teaching of Literature is a research and development center located at the University at Albany, State University of New York. The Center was established in 1987 with funds from the U.S. Department of Education, Office of Educational Research and Improvement, and from the National Endowment for the Arts. The Center's mission is to improve the teaching of content knowledge and critical-thinking strategies that contribute to literary understanding, particularly at the middle and high school levels.

Center-sponsored research falls into three broad areas: 1) surveys of current practice in the teaching of literature, including studies of both what is taught and how it is taught; 2) studies of alternative approaches to instruction and their effects on students' knowledge of literature and critical-thinking abilities; and 3) studies of alternative approaches to the assessment of literature achievement, including both classroom-based and larger-scale approaches to testing.

The Center also promotes good practice in the teaching of literature through conferences and seminars, through the development of computerised bibliographies on research and practice in the teaching of literature, and through publications that present the Center's own research and provide other resources for research and practice. To receive a list of current publications, please write to CLTL, School of Education, University at Albany, 1400 Washington Avenue, Albany, NY 12222.

Director:	Arthur N. Applebee	Center Faculty:	Lil Brannon
			Eugene Garber
Co-Directors:	Judith A. Langer		Peter Johnston
	Alan C. Purves		C.H. Knoblauch
			James Marshall
Assistant Director:	Genevieve T. Bronk		Arnulfo Ramirez
			Wayne Ross
			Sean Walmsley

Abstract

A census of assessment programs in each of the states and the District of Columbia reveals that 61% of the systems assess literature as a part of an assessment of reading. The main focus of these assessment programs is on comprehension rather than cultural literacy. A content analysis of all available proprietary testing programs affecting secondary school students reveals that the majority of items deal with comprehension of the content of the text rather than any literary aspects such as language, structure, or voice. The tests in commercial anthology series also focus on comprehension with more attention to theme and symbol than occurs in the proprietary tests. The one exception to this overall picture occurs in university placement examinations, which contain a high number of items calling for knowledge of literature and literary skills. Essay examination assignments also ask for more complex cognitive strategies than do the objective measures, but these are limited to a few tests for specialized students or to anthologies where they are often supplementary to objective measures.

Contents

Background: The State of Assessment 1

Section 1. Census of State Assessments 2
Purpose and Methodology 2
Results 3

Section 2. Content Analysis of Proprietary Tests 6
Purpose and Sample 6
Analytic Framework for Examining Tests 7
Results 9
- Commercial Achievement Batteries 11
- Commercial Reading Tests 11
- Commercial Literature Tests at the High School Level 15
- College Admissions Tests 15
- College Placement Examinations in Literature 18
- High School Anthology Series 18
- Basal Readers 23

Section 3. Essay Items on all Literature Tests 26
Sample and Methodology 26
Results 27

Section 4. Summary and Conclusions 29

References 31

Appendices 33
1. Interview Schedule 33
2. List of Commercial Tests 36
3. Classification Scheme for Test Items 38
4. Appendix Tables 44

The Current State of Assessment in Literature

Pamela Brody
Carol DeMilo
Alan C. Purves

University at Albany
State University of New York

This report presents the results of a census of state assessments within the United States and of a content analysis of tests produced by proprietary testing companies and textbook producers. It seeks to portray how literature achievement is currently assessed in the United States. The census and content analysis were undertaken from January to June 1988. Literature in this report is defined as poetry, drama, prose fiction, and prose non-fiction that has a belles-lettristic cast to it (e.g., literary biography, the personal and critical essay). We are also defining as literary questions that focus on structure, style, voice, mood, or tone as opposed to those that focus upon the manifest content or message of the text. A fuller discussion of these definitions is to be found in Section 2 of this report.

Background: The State of Assessment

When the International Association for the Evaluation of Educational Achievement (IEA) assessed student performance in the reading of literature over 15 years ago, American secondary students on average scored lower than did the students of several other nations. New Zealand students in particular exhibited a high level of performance (Purves et al., 1973). A subsequent follow-up of New Zealand schools showed that this higher level of performance could in part be accounted for by the fact that the examination system contained a large number of questions that asked students to read literature critically (Purves, 1979b). Students were required to write both about previously studied works and about works that were fresh to the student at the time of the examination.

In the United States, on the other hand, the major examinations such as the College Entrance Examination Board's Scholastic Aptitude Test, the American College Testing Program set of examinations, standardized achievement test programs and existing minimal competence and state-assessment programs, contained no literature section. They might contain an occasional literary passage within a reading comprehension test (Grommon, 1976). Only the College Board's Achievement Test in Literature and Advanced Placement Examination in Literature measured the critical reading of and writing about literature and these examinations were limited to a select proportion of pre-university students. Although there was some emphasis on literature on the New York Regent's Examination, it, too, was reserved for the college-bound. Thus, examinations of students' critical reading abilities was not an integral part of the schooling experience for most students, and only a peripheral part for a small segment of the college-bound.

While one might not necessarily approve of it, the fact is that the examination system drives the English curriculum (Langer & Applebee, 1987); the absence of literature in university entrance examinations and large-scale assessments means that literature, and particularly critical understanding of complex texts, will get short shrift.

There is little information as to what changes, if any, in the assessment practices of various institutions or corporations have occurred as a result of the educational reform movement of the 1980s. This movement should have had an impact on literature instruction and assessment, but in fact, there has been no comprehensive study of testing practices in literature achievement since the 1976 survey by Grommon. The recent writing of E.D. Hirsch, Jr. (1983, 1987), in particular, would suggest that educational reform should include an emphasis on knowledge, much of it literary knowledge, and that such emphasis should be reflected in tests. Has this in fact occurred? Discussions of the reforms suggest that there is little attention to literature in assessments of reading or writing, primarily because these assessments are seen to deal with "functional literacy." In writing assessments, for example, the tasks could conceivably include both tasks which ask students to write about literature and tasks that encourage them to produce literary texts. Such is the case with the National Assessment of Educational Progress and the new California Assessment (Applebee, Langer, & Mullis, 1986; Cooper, 1987). But how extensively do the testing programs encourage expressive writing concerning the students' responses or writing that taps their creative imagination? In reading assessments, as well, it is possible for literature to be included, but we do not know the extent to which this occurs or the nature of the questions. In light of such issues, this study seeks to provide a broad overview of the current situation. Section 1 of this report examines state-mandated assessments. Section 2 examines proprietary tests of all kinds, including commercial achievement batteries, literature tests, college admission tests, college placement tests, tests provided as part of high school literature anthologies, and tests accompanying grade 7 basal readers. Section 3 discusses the essay items that were included as a part of some of these tests. Finally, section 4 provides a summary and discussion of the results as a whole.

Section 1
Census of State Assessments

Purpose and Methodology

The purpose of the census of state assessments was to provide an overview of state-mandated testing programs and their use of literature. The tests that are discussed in Section 2 may be selected by a school, a teacher, or a student (if applying to a particular institution), but only at the state level might one see a consistent policy about the measurement of achievement in a given area. The state census was based on a telephone interview of testing directors for each of the 50 states and the District of Columbia, all of whom responded. The interview schedule was developed with the cooperation of the New York State Department of Education testing division and based on a questionnaire developed by the Chief State School Officers internal information group. The questions were standardized and appear as Appendix 1. The interviewers did not try to lead the respondents; thus, the results reflect the respondents' interpretations of the language of the interview schedule. Some states requested that the detailed results be confidential so that reporting is general and by region, rather than state by state.

Results

Only two states stated they had no statewide testing or assessment program and no plans for starting one. All of the other states have or are planning some sort of a testing program. In most states (41), the tests or assessments are administered annually or more frequently and over half (26) develop their own tests exclusively.

Table 1 shows the percentage of states that test student knowledge of literature or ability to read and write about literature in some form: as a separate subject, as a part of a reading assessment, or as a part of a writing assessment. The table gives both regional and national totals. It is safe to say that the testing of learning in literature is not emphasized as a separate topic by most states, but as an aspect of reading. What this means in practice is that reading assessments either include some passages from literary works in their mix of sources of texts or include a literature section as a sub-test within a reading assessment. Two states have a humanities assessment and include literature as an aspect of general cultural and intellectual history. The regional data show that states in the Northeast and West develop their own tests somewhat more than do states in other areas.

When states do test literature, what aspects of literature are assessed? A summary of the responses of those states which test literature either separately or as part of a reading or writing assessment is presented in Table 2, which indicates what percentage of states assess certain aspects of the subject literature. It is clear, as we suggested above, that a minority of states (mostly in the South) use literary texts but do not test on specific aspects of the subject, literature. Most states that include literature claim to include items that deal with the texts as literary artifacts and to cover such literary aspects as language, structure or genre, and voice, tone or mood. Fewer than a quarter of the states measure student knowledge of specific authors and titles (these mostly in the Northeast), literary terminology, or general cultural information. In addition, three state assessments include items on the interests, attitudes, or attributes of readers.

The assessments are used mostly for curriculum revision and for placement or remediation, and to a lesser extent for promotion. Given the limitations of the interview format, we cannot tell what curriculum revisions have resulted or are impending. Since the tests are primarily reading tests, literature serves the most common functions of reading comprehension assessment generally: placement, retention, and graduation.

Because many states refused to allow us to examine the actual tests, it is difficult to verify whether the claims by the state assessment groups are matched by the actual items employed, except where a state used a commercial or proprietary test, which we shall examine in the next section.

Table 1
State Assessments of Literature Achievement
1987-1988 School Year

How Literature Is Assessed	Number of States				
	Northeast (n=11)	Southeast (n=13)	Central (n=12)	West (n=15)	Total (n=51)
As a Separate Area					
Commercial Test	0	0	0	0	0
State Developed	3	0	1	1	5
Both	0	0	0	0	0
Through Reading					
Commercial Test	3	1	4	1	9
State Developed	5	1	1	3	10
Both	0	8	1	3	12
Through Writing					
Commercial Test	0	0	0	0	0
State Developed	1	0	1	1	3
Both	0	0	0	0	0
Through Humanities					
Commercial Test	0	0	0	0	0
State Developed	2	0	0	0	2
Both	0	0	0	0	0
Total Assessing					
Literature in Any Way	8	10	6	7	31
Percent of States	73%	77%	50%	47%	61%

Sample was the 50 U.S. States and Washington, D.C.

Table 2
Characteristics of Literature Assessment in State Assessment Programs

	Number of States				
	Northeast	Southeast	Central	West	Total
Aspects of Literature Assessed					
Comprehension Only	0	6	1	2	9
Reading Comprehension and Literature Content	6	3	2	4	15
Specific Authors/Titles	3	0	1	0	4
Literary Terms	4	3	2	4	13
Language	4	3	2	2	11
Structure or Genre	6	3	2	3	14
Voice, Tone, or Mood	3	3	2	4	12
Folklore/Cultural Information	3	0	1	2	6
Readers' Attributes (e.g., attitudes & interests)	1	0	2	0	3
Item Type					
Multiple Choice	6	7	3	5	21
Essay	2	4	0	2	8
Test Scoring Type					
Norm-Referenced	2	4	0	0	6
Criterion-Referenced	4	7	3	5	19
Uses					
Graduation	3	4	0	2	9
Placement/Remediation	2	5	1	2	10
Curriculum Revision	5	2	2	2	12
Award	1	0	0	0	1
Promotion/Retention	2	2	0	0	4
Allocation of Funds	1	2	0	1	4
Accountablity	0	2	1	0	4
Score Reporting					
Pupil	5	8	3	6	22
Class	4	6	3	4	17
School	6	8	2	6	22
District	6	8	1	6	21
State	6	6	2	5	19
Number of States	6	9	3	6	24

In summary, the majority of the 31 states that test literature in any way treat literature as an adjunct to or aspect of reading. Fewer than half appear to focus on the literary or aesthetic aspects of the literary text. In a minority of these states are such matters as cultural or background information, specific authors or texts, and literary terminology and theory tested. Since only two of the states report that these particular measures are used to help determine promotion or graduation, it would seem that Hirsch's (1983, 1987) recommendations to use this type of examination to help insure cultural literacy are not as yet implemented. Whether implementation will occur remains to be seen. Since two thirds of these states that test literature use commercial tests exclusively or in combination with state developed ones, an analysis of these tests should give a strong indication of the focus of assessment in literature at the state level. Section 2 provides such an analysis.

Section 2
Content Analysis of Proprietary Tests

Purpose and Sample

This aspect of the study sought to gain a closer look at testing practices with respect to literature achievement as found in the published tests available to teachers, school districts, and states. The tests studied included those prepared by both non-profit and for-profit publishers who sell their tests to teachers, school districts, states, or institutions of higher education. The selection of tests was made using Tests in Print; all tests that could possibly have a literature section were selected for analysis. Excluded were tests on single works. Commercial tests (n=16) included three types:

1. Commercial achievement batteries, which usually include mathematics, science, social science, language arts, and a separate reading segment (n=12 with a total of 29 forms since some batteries were multilevel).

2. High-school level literature tests (n=2).

3. Reading tests which may have literature sub-scales, but which purport to measure reading ability or achievement (n=2).

There are two other testing program that were considered:

4. University entrance, and

5. Placement programs controlled by the American College Testing Program (ACT) and the College Entrance Examination Board (CEEB).

From these were selected:

(a) Preliminary ACT, ACT, Scholastic Aptitude Test (SAT), Preliminary Scholastic Aptitude Test (PSAT), and Advanced Placement tests.

(b) College-Level Placement Tests in Literature (CLEP).

Of these placement programs, only the Preliminary ACT, the SAT (a sample test only), Advanced Placement (essay segment only), and five CLEP placement tests were released by the organizations for review.

Two other groups of tests were analyzed:

6. Those tests that appear in the Teacher's Guide or testing program of the five major publishers of secondary literature series as indicated by sales figures. Since each series spans grades 7 through 12 but is generally consistent in its approach within a grade level (i.e., having the same assessment format), and since each year consists of some 12 to 20 units, each of which has a test, we sampled from grades 7, 9, and 11 in each series and selected every other unit.

7. Tests accompanying major basal series in Grade 7. Because the anthology series begin with Grade 7, where there may be a competing basal reading text which also contains a number of literary selections, we decided to include the five major basal series texts and tests for Grade 7. In this way, it is possible to get some sense of the alternate forms of testing confronting a student in the seventh grade.

All the tests analyzed are listed in Appendix 2. For reasons of confidentiality and because the purpose of the study is to determine trends not to describe specific publishers, the analysis does not include references to specific tests except as a category (e.g., College Placement).

Analytic Framework for Examining Tests

Perhaps the most widely used framework for the evaluation of learning in literature at any level is that developed in the 1960s for the International Association for the Evaluation of Education Achievement (IEA) study of literature education in ten countries (Purves, 1973). It depicts the domain of literature education in terms of content and behaviors. That framework was presented in an elaborated form by Purves in Bloom, Hastings, and Madaus (1971).

With the cross-classifications of the framework, any curriculum can be mapped and from it the specifications for assessment and evaluation can be written. A modified version of that framework specifying how the approach can be made more specific to the measurement of the reading of a single text was developed by Cooper and Purves (1973; reprinted in Purves, 1979b). This study used a combined version of the two (Figure 1).

The first set of classifications deals with the broad area of content: author and title; literary terms; comprehension of content, character, theme, language, structure, voice; and knowledge of general cultural information. The second set of classifications deals with the cognitive demand of the item: recall, describe, apply, analyze, interpret and generalize. A set of sample items for each classification is presented in Appendix 3. Only items which refer to

Figure 1
Framework for Analyzing Test Items in Literature

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title							
Literary Terms							
Comprehension/ Inference							
Characterization							
Theme/Moral							
Language							
Structure							
Voice							
Cultural Information							
Reader's Attributes							

literary passages are included in the data analysis; the literature series and basal readers may also have vocabulary and language tests not directly related to the literature selection. Items in the tests include matching, fill-in-the-blank, short answer, and multiple choice. The clear majority are multiple choice.

After training, two raters coded the items on each test according to the grid, and reliability checks were made on 10% of the ratings with agreement on over 95% of the verifications. For each test, the number of items in a category was converted into a percentage of the total items in the test, and then the percentages were averaged over the type of test. For the multi-level batteries, averages over different levels of the same test for each of the 12 batteries were calculated. Overall achievement battery means were then figured by calculating the mean over the 12 published tests. In each anthology series, there were tests to accompany the completion of a unit. There were 6 to 10 units included in an anthology volume, with the number of items per unit test ranging from 15 to 95. Because of the large number of tests to be analyzed and because the tests remained consistent from unit to unit within a series, only every other unit test was coded.

Results

Figure 2 presents an overview of the content of literature items in the proprietary tests, those tests which exist independent of a particular literature program; a corresponding set of detailed tables can be found in Appendix 4. It is clear that the focus of these tests is on the comprehension of content, not unlike the major emphasis of the state assessments, since many of the latter use one of these tests. Test items focus on the texts in terms of their manifest content. Only in the CLEP program is there more than a minimal emphasis on knowledge, primarily of authors and titles. The reason for this emphasis may be that only at the college placement level is there some expectation that a body of knowledge is a part of the literature testing program. That expectation does not appear in the college entrance group, where the tests focus on textual reading and analysis. There is relatively scant attention paid to aspects of the text other than content, although the reading tests and the college admissions tests contain some items dealing with characters. Notably absent from the tests are any substantial number of items dealing with such characteristics of literature texts as language, structure, and voice.

A similar pattern emerges in the cognitive demand implied by the items in these proprietary tests. Most items in all tests focus on recognition and recall and on the application of knowledge to the given text (Figure 3). There is relatively little attempt to deal with such complex mental operations as analysis, interpretation, and generalization. Before one becomes too critical of the tests on this account, one must recognize that such operations are difficult to measure using a multiple-choice format, but are more conducive to essay examination or at least to extended short answer.

Within most types of commercial test, the variation of content and cognitive demand between different publishers' tests was small. (The exception to this was the achievement batteries.) It would appear that within most categories of tests there exists a general pattern, perhaps dictated by market forces. The next few pages will take up each category of test to determine its specific nature.

Figure 2
Content of Literature Items in Proprietary Tests

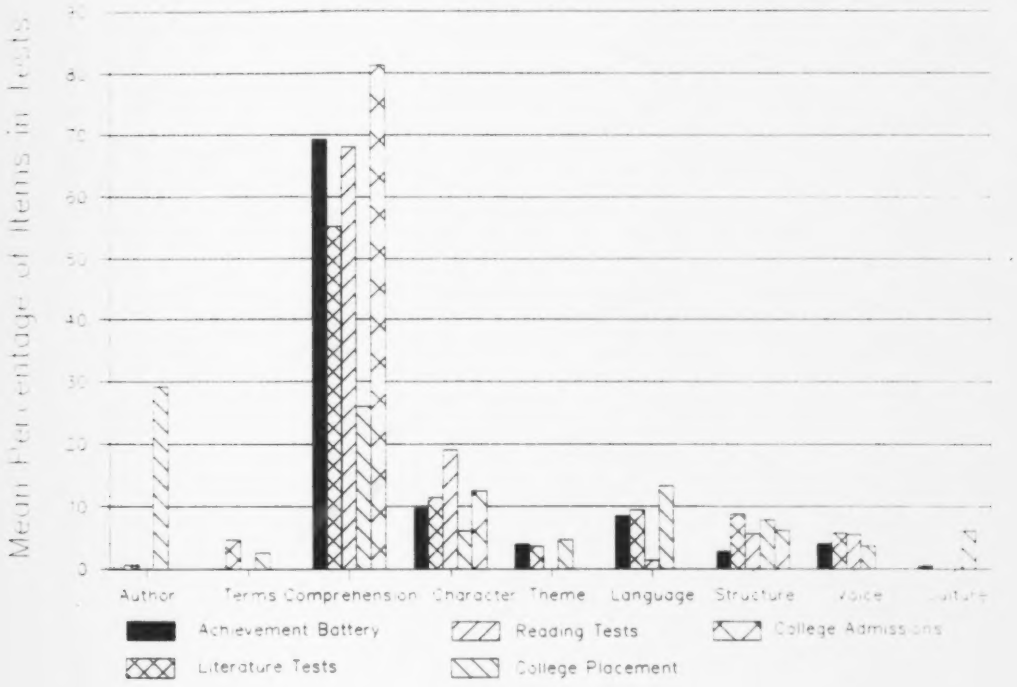
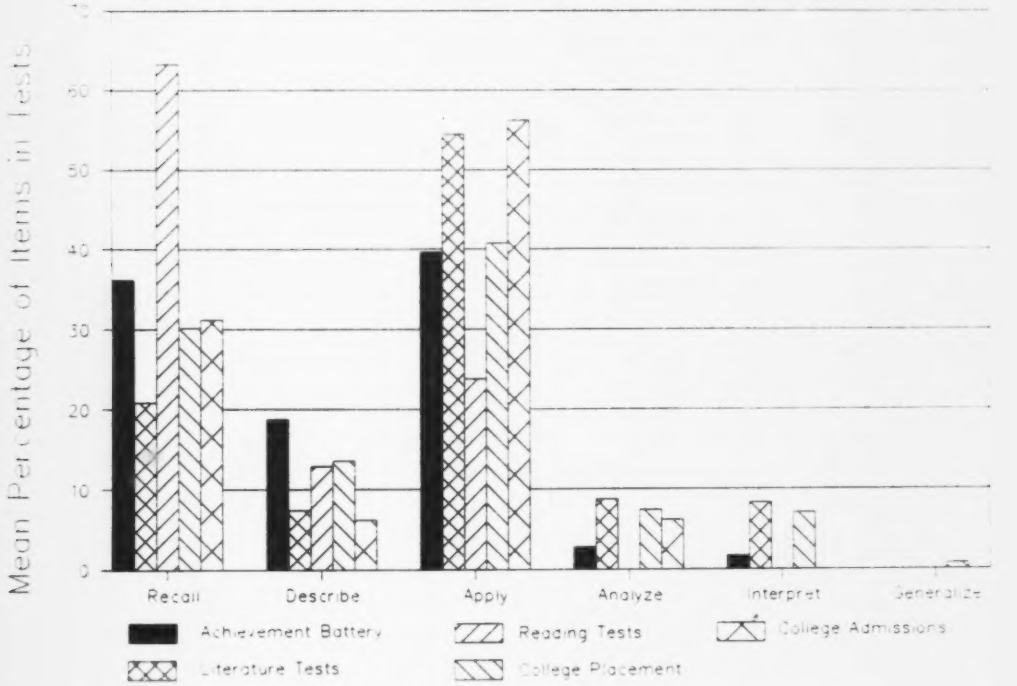


Figure 3
Cognitive Demand of Literature Items In Proprietary Tests



Commercial Achievement Batteries

The content of most of the items in the commercial achievement batteries is either comprehension or drawing inferences (69%) (see Figure 4). Of the remaining content categories, only characterization and language compose 5% or more of the items on the average test. Most publishers include one or two theme, character, structure, voice, tone, or mood items in each reading subtest. Because 75% of the batteries include a poetry selection, about half of the items in the language category involve determining the meaning of figurative language or applying knowledge about diction and metrics to poetry. The items in the structure category most often deal with the organization of the passage, with conflict, or with genre. Out of all items in the 29 tests, only between one and three items were found which measure knowledge or understanding of specific authors, literary terms, or known themes or topics.

These reading subtests are not homogeneous, however. The number of literature items per test ranges from 4 to 39. The percentage of literature items in the reading comprehension subtests ranges from 0 to 76%, with an average subtest utilizing literature in 44% of the items. Some publishers place a strong emphasis on literary questions, and others have few or no formal literary items. Table 3 shows how the content was distributed. The only items included (as with the rest of the study) are those based on literature or which use a literary passage as a prompt.

The majority of publishers have one or more items within each of the following literature categories: characterization, theme, language, structure, and voice. The category most often totally omitted is structure (omitted from 33% of the tests); language and theme are omitted in 25% of the achievement tests. The level of difficulty or cognitive demand (see Figure 5) also has significant variability. Table 4 shows how cognitive demand is distributed among publishers.

These commercial achievement batteries clearly do measure reading comprehension, but only superficially measure literary ability or knowledge. When items do cover literature objectives, there is generally little repetition of the content to improve the reliability of measuring the specific objective. The variation among these batteries suggests that one cannot assume that any two tests are equivalent in their coverage of literature or in their level of cognitive demand.

Commercial Reading Tests

Of the ten commonly used commercial reading tests selected for review, only two contain any literary passages (passages that could clearly be identified as coming from poetry, drama, prose fiction, the literary essay or the literary biography). On these tests, an average of 37% of the reading comprehension items refer to the literature passages. Most of the content covered by the items (Figure 6) is either comprehension or inference (68%), or characterization (19%). In the characterization category, 16% of the items involve recall, which also might be considered comprehension (see Appendix Table 2). There are some items about voice (6%) which could be considered literature questions instead of pure comprehension. The cognitive demand is at the recall (63%), describe (13%) or apply (24%) levels. In general, reading tests are more homogeneous than achievement batteries but cover literature content infrequently.

Figure 4
Content of Literature Items in
Commercial Achievement Batteries

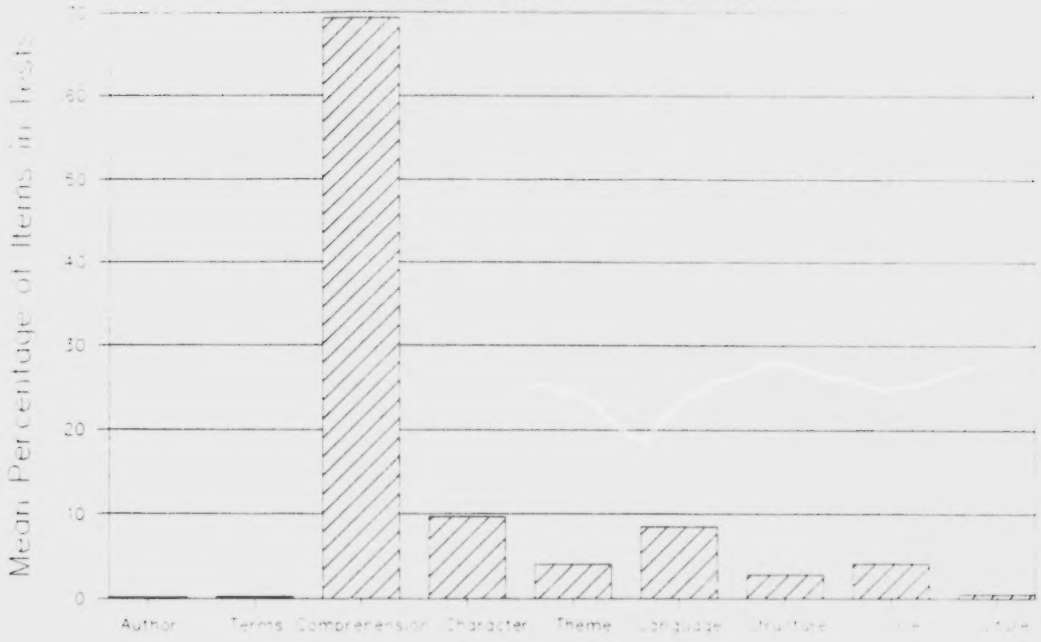


Figure 5
Cognitive Demand of Literature Items
In Commercial Achievement Batteries

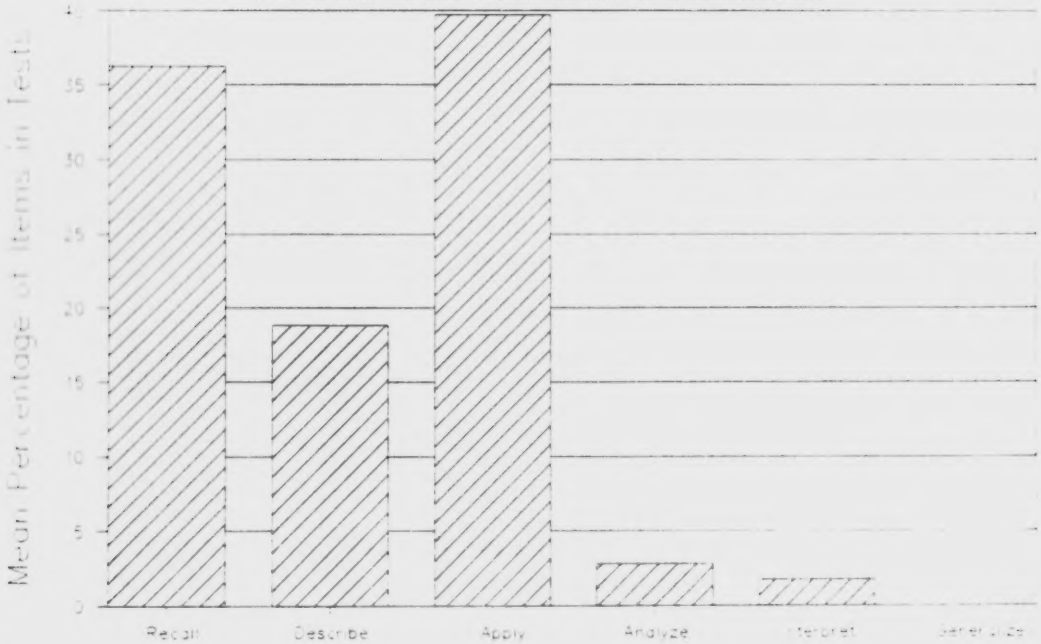


Table 3

Variability of Content in Achievement Batteries

Number of tests at various percentage levels of items with literature content

Comprehension Only	.1-9.9%	10-19.9%	20-29.9%	30-39.9%	40-49.9%	Over 50%
1	0	4	2	3	1	1

n = 12 tests

Table 4

Variability of Cognitive Demand in Achievement Batteries

Number of tests at various percentage levels of items with given cognitive demand

	60% or more	45-59%	30-44%	15-29%	less than 15%	none
Recall	2	3	1	4	1	1
Describe/ Summarize	0	1	0	7	4	0
Apply	1	4	4	2	1	0
Analyze	0	0	0	0	8	4
Interpret	0	0	0	0	7	5

n = 12 tests

Figure 6
Content of Literature Items in
Commercial Reading Tests

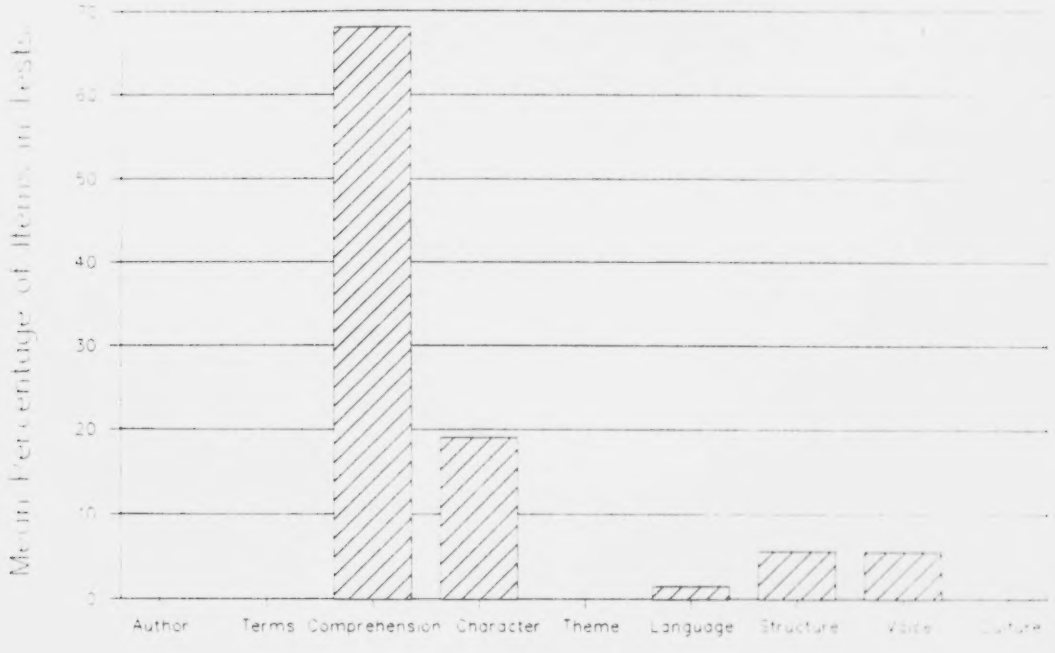
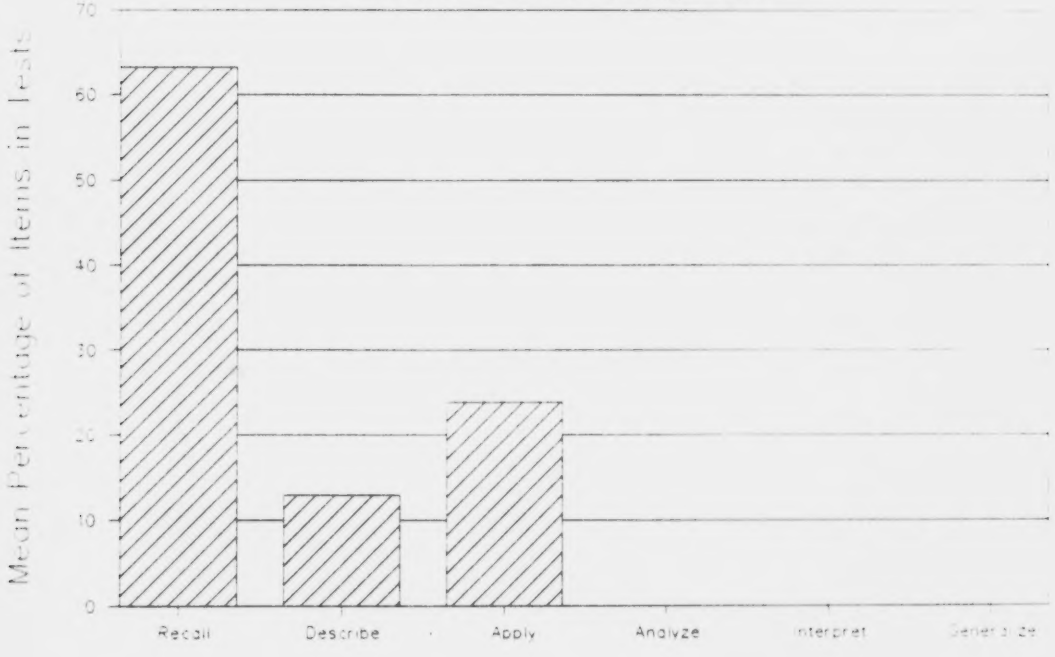


Figure 7
Cognitive Demand of Literature Items
In Commercial Reading Tests



Commercial Literature Tests at the High School Level

Only two commercial tests specifically designed to measure interpretation and knowledge of literature were found for secondary students. These commercial tests are also multiple choice tests. Most items are based on literature passages that are given as part of the test.

Although these tests were designed as measures of literature achievement, the largest number of questions on the test fall into the comprehension and inference category (55%). Understanding symbols is a special form of inference and these items formed a visible part of the tests; they were at least twice as common in these literature subtests as in the reading comprehension tests and achievement batteries. (See Appendix Table 3 where this difference can be seen in the intersection of "analyze and interpret" with the comprehension category (nearly 7% of these items as opposed to 1 or 2% in the other tests).) There is also greater emphasis on literary terms, language, structure and voice than in the achievement batteries (see Figure 8).

The cognitive demands elicited by these literature tests are similar to those of the college admissions exams and appear "higher" than those elicited by the achievement batteries, reading tests, and college placement exams (see Figure 9). The majority of the items (54%) required application of terms. These tests also have the highest percentage of items at the level of analysis (9%) and interpretation (8%).

One may conclude that in content these tests are similar to achievement batteries and reading tests although with a somewhat more diverse range of cognitive demand. They do not appear to be what many might term literature tests, however.

One publisher has a unique type of literature assessment available for specific literary works. The tests are in the format of item pools for use on microcomputers. So far only two tests of specific plays are available. The educator can select from 246 multiple choice or true/false items, 20 short answer items, and 10 essay questions. The majority of the objective items are comprehension/inference (46%) and characterization (42%). The cognitive demand is 62% recall and 37% application in those items. The short answer items require higher cognitive behaviors, with 50% of those items requiring interpretation. The essays generally asked the student to present a persuasive argument. The actual content and cognitive demand of a given test produced from the item pool depends on the specific items selected by the teacher.

College Admission Tests

Two test programs used either as a criterion for college admissions or as a pre-admission scholarship test were included in this part of the study. In these tests, literature was used as a tool for measuring reading comprehension (see Figure 10); therefore, the majority of the test items fall in the comprehension/inference content category (81%). This is the highest percentage of inference among the test types included in the study. The other content categories included were characterization (13%) and analysis of structure (7%).

Figure 8
Content of Literature Items in
High School Level Literature Tests

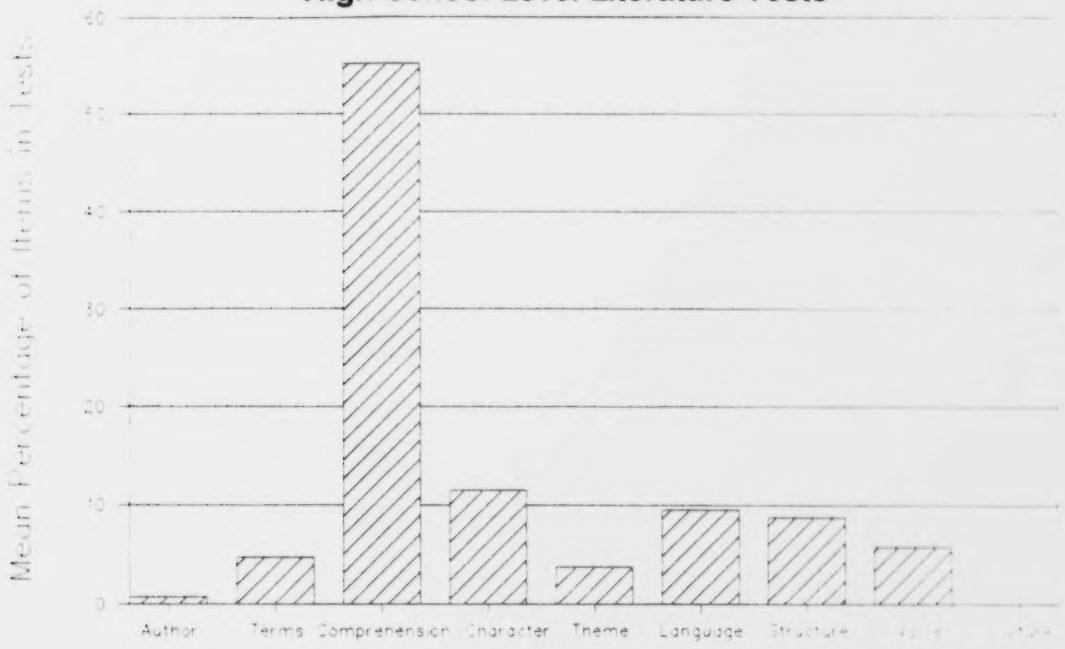


Figure 9
Cognitive Demand of Literature Items
In High School Level Literature Tests

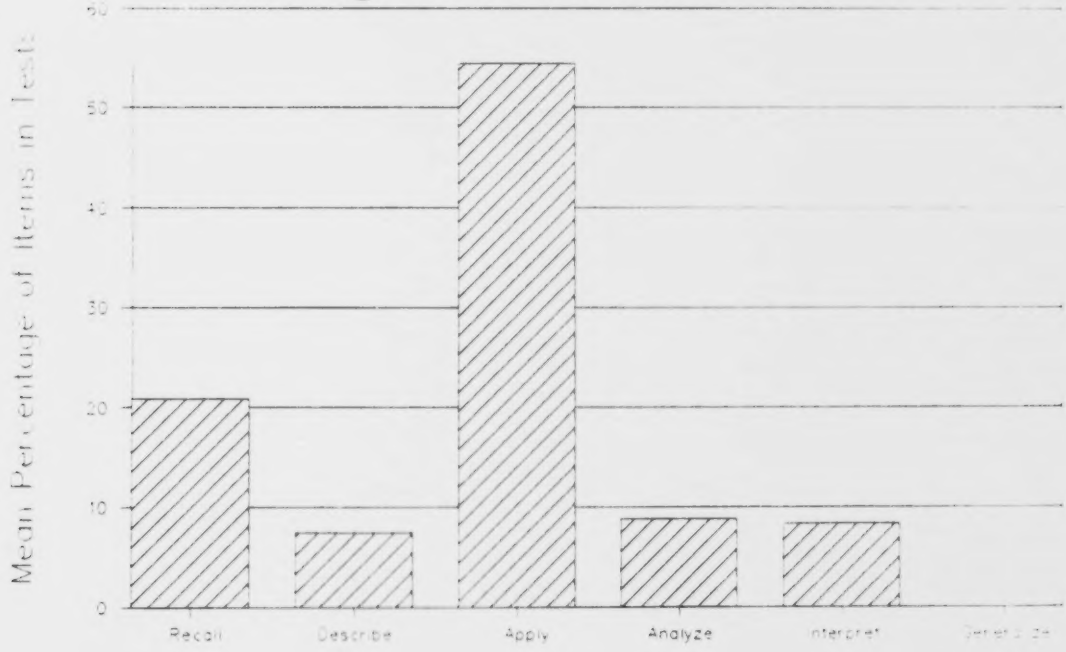


Figure 10
Content of Literature Items in
College Admissions Exams

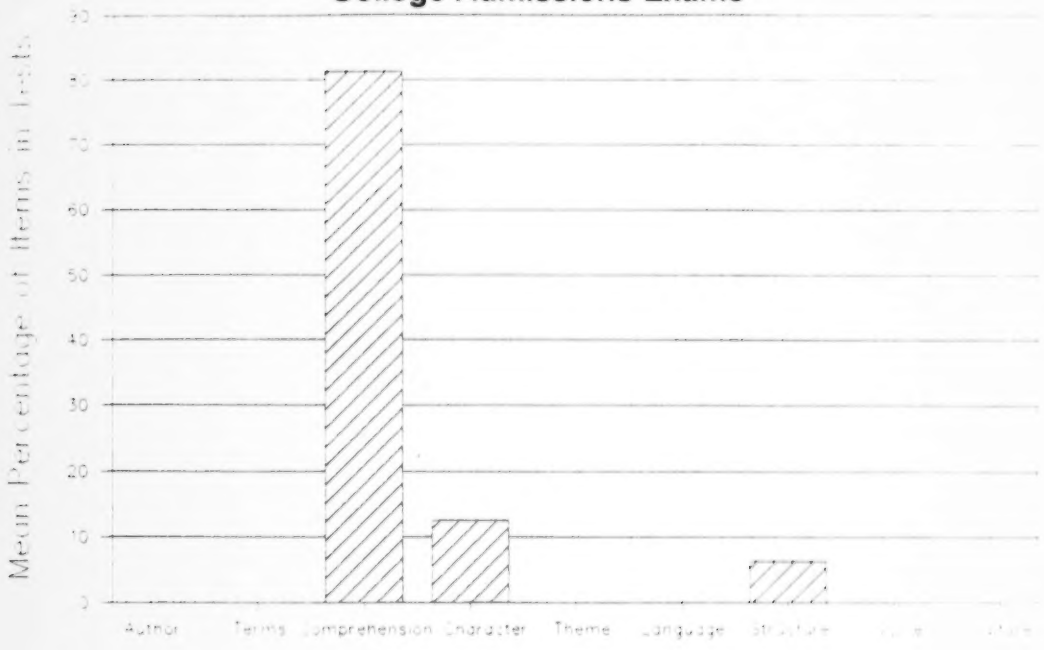
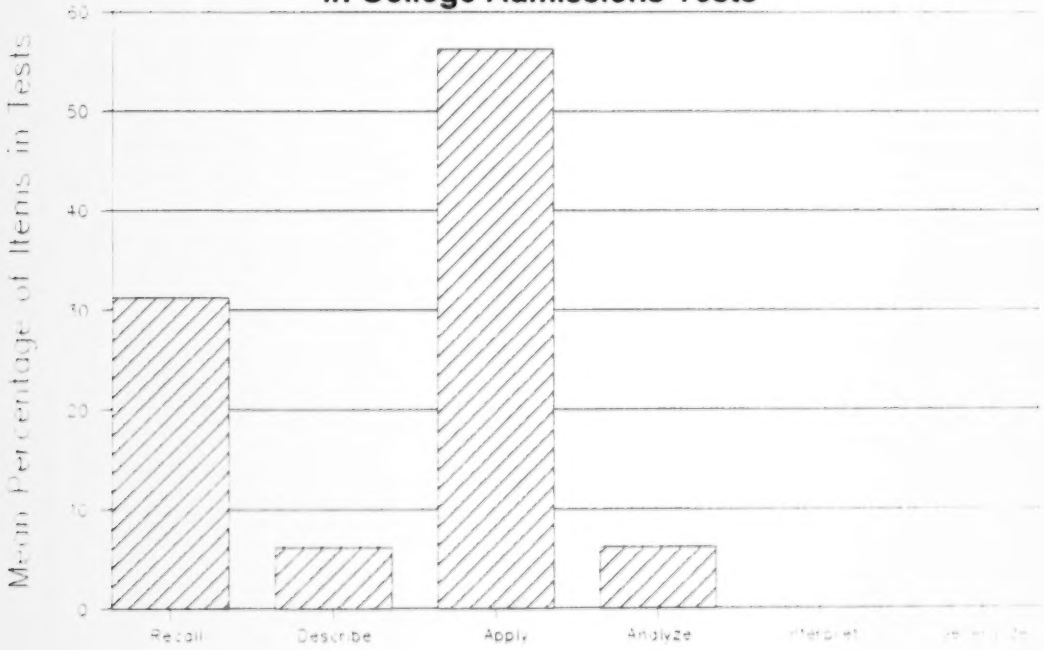


Figure 11
Cognitive Demand of Literature Items
In College Admissions Tests



The most common cognitive level of the items requires application of knowledge (56%). The "highest" level of cognitive demand tested is analysis; all the other test categories except the reading tests required some interpretation (see Figure 11).

Again these tests can best be seen as reading tests that include some literary passages but do not treat them distinctively.

College Placement Examinations in Literature

These examinations are used to place incoming college students in the appropriate level of a literature course, or to award college credits. Five examinations were analyzed, four of which had 100% literature content. The fifth, an exam in humanities, contained 50% literature items.

The largest content category in the college placement examinations is specific authorship or title of literary works (29%) (see Figure 12), with most of these items requiring recall or recognition (20% of total test). None of the other test groups studied have a significant percentage of items in this author/title category. The next largest content category is comprehension/inference (26%). As in the high school level literature tests, a portion of this category includes identifying and interpreting symbols, inflating the number of items devoted strictly to comprehension.

The college placement examinations also have a high percentage of items in the content areas of language (13%), mostly application of knowledge of figurative language and application and analysis of metrics; structure (8%), mostly application and analysis of organization and genre; and cultural information (6%), mostly recall and application of knowledge about Biblical and mythological texts or known themes. There are smaller numbers of items involving literary terms, characterization, theme, and voice.

The cognitive demands required by the items in this test type are higher than those of the achievement batteries and reading tests (see Figure 13), but because of the emphasis on authors and cultural knowledge, lower than the demands of the high school level literature exams, where 40% of the items on a test demanded application and 30% recognition. The college placement exams are the only group with a measurable amount of items requiring generalization of knowledge (1%).

Of all the tests studied, this group has the broadest coverage of the domain of literature and to a lesser extent, of cognitive tasks. This is the one set of tests that assumes a core of knowledge on the part of the student.

High School Anthology Series

In addition to the proprietary tests and entrance and placement examinations, there is another body of readily available tests of literature: those that occur within the anthology series. These, perhaps, are the most ubiquitous in the schools, but it is hard to be sure since we do not know the extent to which they are used. The anthologies are usually divided into units, and each unit has some sort of test, normally an objective measure and an essay assignment.

Figure 12
Content of Literature Items in
College Placement Exams in Literature

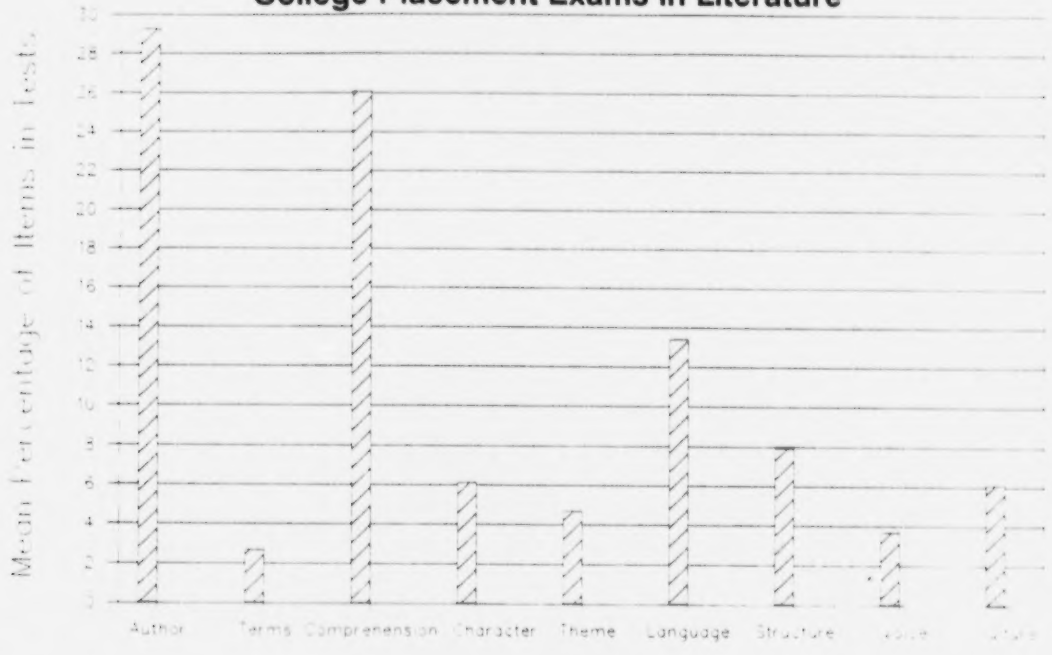
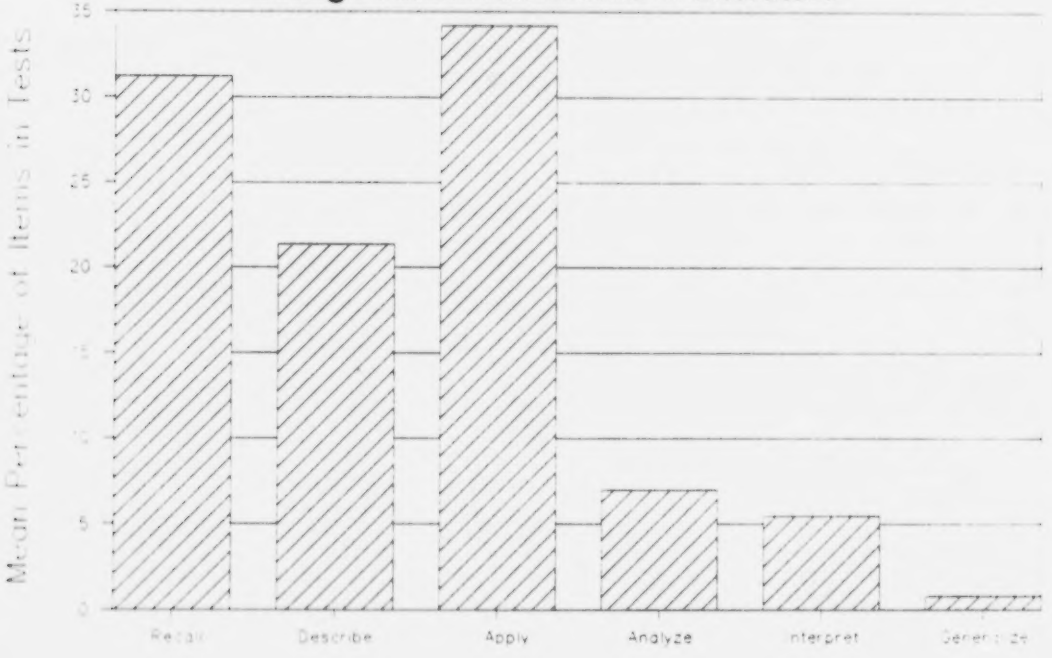


Figure 13
Cognitive Demand of Literature Items in
College Placement Exams in Literature



In the area of content, most publishers emphasized comprehension. However, there is a great deal of variability among anthology series with regard to the literature items on the tests. The range for language items is .3% to 9%. For structure, the range is .5% to 4%. The range in the area of characterization is 8% to 46%. In the area of literary terms the range is .3% to 30%.

There was also variability among the units of an anthology series across grades. For example, there is a range of 7% to 30% for one publisher on literary terms. For a different publisher, there is a range of 16% to 46% in the area of characterization. Interestingly enough, the increase in percentage does not correlate with advancing grade level within the series. There is variability in the levels of cognitive demand tested in the items across publishers, but not in the areas of interpret and describe, as Table 5 shows.

One explanation for the variability found both across and within publishers is that most publishers put together a team of testmakers who are independent of the authors. Usually the tests are not developed by the people who work on the literature series, and the job of test writing is contracted out so that more than one person may write the tests for the various grades within a series.

Despite this variability within and between publishers, it is useful to examine trends over the course of a student's progress through secondary school. The discussion that follows focuses on between grade comparisons. Figure 14 illustrates the differences in content representation between grades.

Most of the anthology test questions deal with knowledge of the readings in the textbook which students had completed prior to taking the test. The rest of the questions deal with knowledge of vocabulary presented in the context of their prior reading. The content category of most of the items is comprehension/inference for each of the grades (Grade 7: 46%; Grade 9: 44%; Grade 11: 34%). Questions relating to literature skills consisted of between 0 and 22% of the items, about 22% referring to characterization for Grade 7. The content categories with the fewest items across grades are language and voice, but there is an increase with grade level of items dealing with literary terms, language, structure, and voice. At Grade 11, the distribution of items across content area is more balanced than at Grades 7 and 9, though comprehension dominates even at Grade 11.

Figure 15 summarizes the differences in cognitive demand between grades. Items calling for pure recall make up the overwhelming majority for all three grades (Grade 7: 69%; Grade 9: 51%; Grade 11: 47%). The next most common cognitive demand is apply (Grade 7: 11%; Grade 9: 36%; Grade 11: 22%), but the higher levels of thinking such as analysis and interpretation are called upon less frequently.

It seems clear from this analysis that the anthology tests do not present a clear or logical pattern of progression from grade to grade, either in the areas of content tested or in the cognitive demand of the test items. There is no movement from the simpler to the more complex in terms of cognitive demand, nor is there a deepening or broadening of the literary aspects of reading. Any structure or sequence to the anthology testing program eludes this analysis.

Figure 14
Content of Literature Items in Anthology Tests

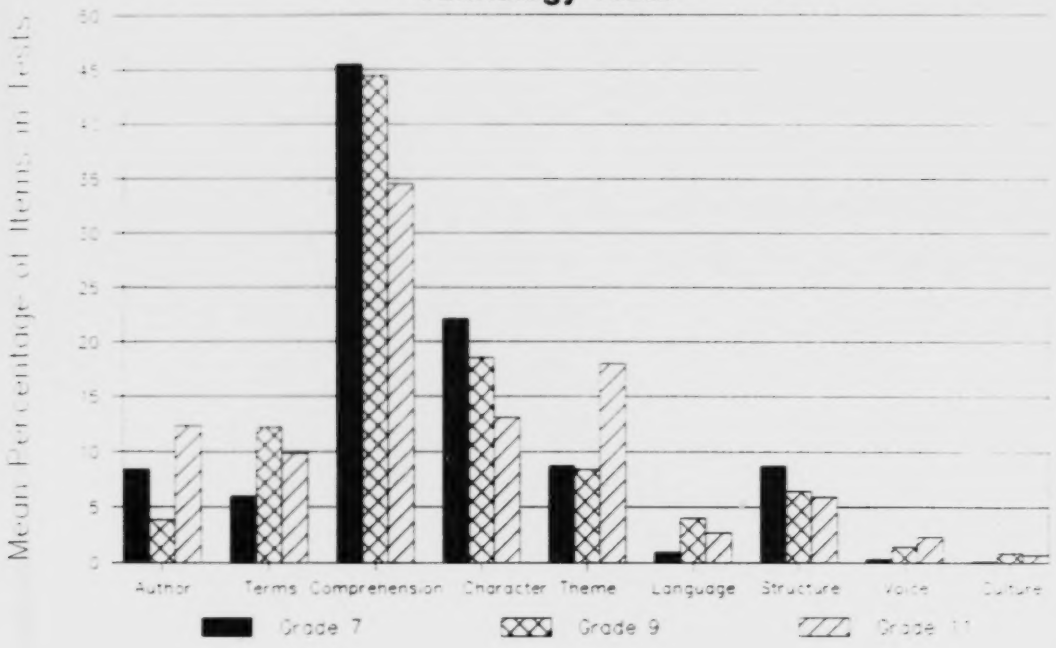


Figure 15
Cognitive Demand of Literature Items In Anthology Tests

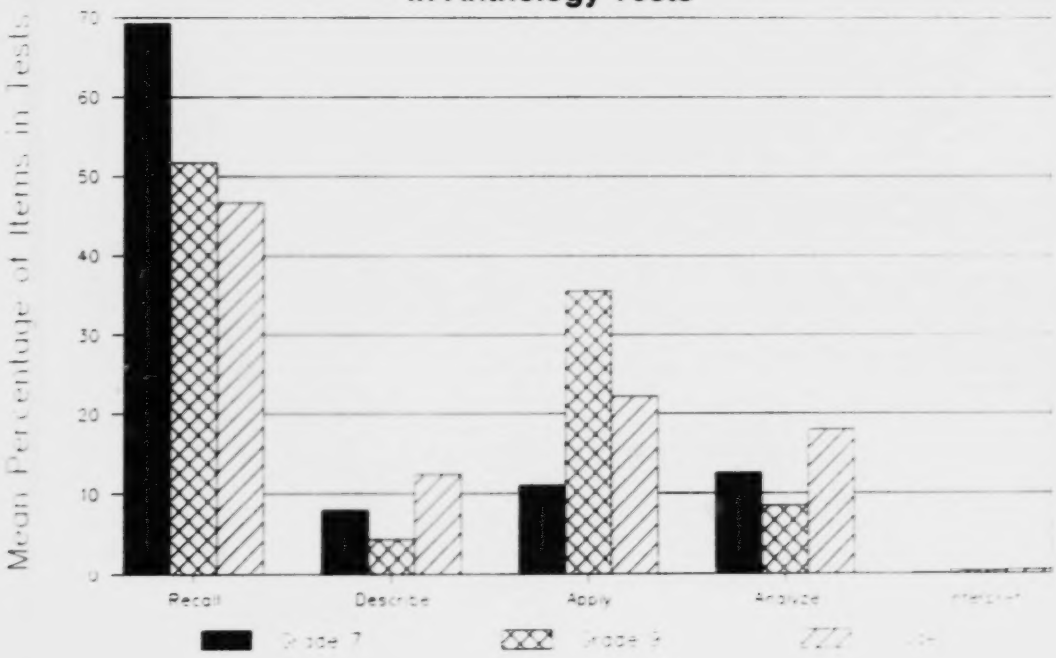


Table 5

Variability of Cognitive Demand in Anthology Tests

Number of publishers at various percentage levels of items with given cognitive demand

	50% or more	35-49%	20-34%	10-19%	1-9%	0-.99%
Recall	8	4	0	2	1	0
Describe	0	0	2	2	8	3
Apply	1	4	3	3	3	1
Analyze	1	1	2	3	4	4
Interpret	0	0	0	0	1	14

n = 15 (5 publishers by 3 grade levels)

Basal Readers

Since many basal reading series have a seventh-grade volume, which may supplement or supplant the literature program, it was important to compare the reading and the literature programs.

There is some variability of content in the basal reader tests. The total proportion of literary items is similar for most of the publisher (12% to 20% of the items), but one publisher has a significantly larger proportion of items other than comprehension (51%). The content categories are relatively uniform with the following exceptions: structure ranges from .9% to 24%, language ranges from 3% to 15%, and characterization ranges from 2% to 8% of the item content.

As might be expected, comprehension questions make up the vast majority of basal reader items (79%). The next largest content areas are structure (6%), language (5%), and characterization (5%). There are almost no questions about specific authors and titles. As in the Grade 7 anthology tests, voice and cultural information items are rarely included in the basal reader tests (see Figure 16).

In the area of cognitive demand, recall items made up 46% and apply items 29% of the basal reader tests. There is a significant percentage of content in the areas of analyze (12%), describe (8%), and interpret (5%). There is some variability between publishers in this area, as is shown in Table 6 .

The multiple-choice items of the basal readers require higher levels of cognitive demand than do the Grade 7 anthologies (see Figure 17).

However, an important difference between the tests accompanying these two types of textbooks is that there are no essay questions included in the basal reader tests. Seventh-grade students in a literature program are thus more likely to get a balanced fare of reading and writing tasks than if they were in a basal program. Their reading tasks, however, would place lower cognitive demands on them. If the multiple choice tests in the anthologies appear to neglect many of the aspects of literature that critics call for, and appear to tap lesser skills, perhaps the essay questions fill the gap.

Figure 16
Content of Literature Items in
Grade 7 Anthologies and Basal Readers

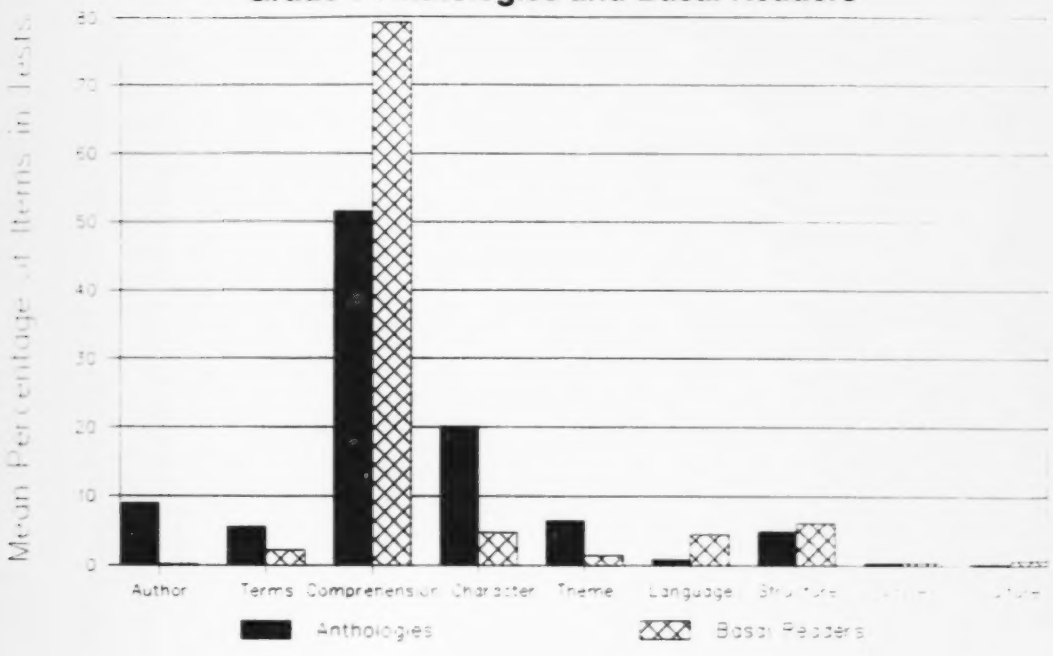


Figure 17
Cognitive Demand of Literature Items in
Grade 7 Anthologies and Basal Readers

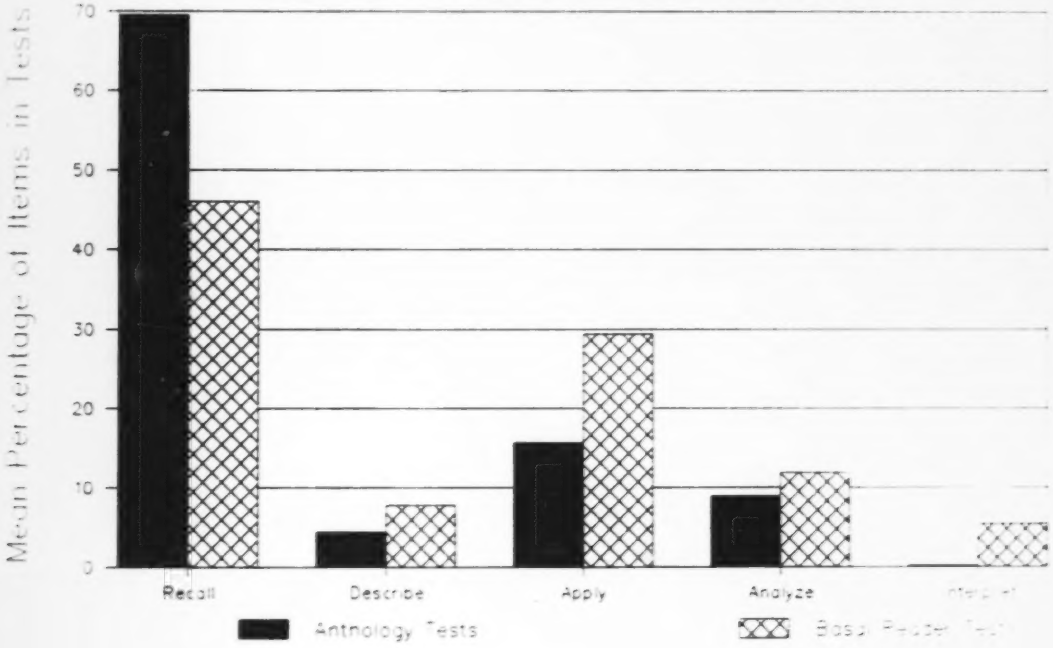


Table 6

Variability of Cognitive Demand in Basal Reader Tests

Number of publishers at various percentage levels of items with given cognitive demand

	50% or more	35-49%	20-34%	10-19%	1-9%	0-.99%
Recall	3	0	2	0	0	0
Describe	0	0	0	1	4	0
Apply	0	1	4	0	0	0
Analyze	0	0	1	1	2	1
Interpret	0	0	0	1	4	0

n = 5

Section 3
Essay Items on all Literature Tests

Sample and Methodology

Next was conducted a content analysis of the essay questions in the anthologies, one college placement test, two state tests (one a graduation test), and a statewide assessment in writing that used literature among its topics. This sample constitutes the available pool of published essay examinations. However, this pool may not represent the total range of topics and assignments that students are given because many teachers create their own assignments or develop school- or district-wide tests. What is represented here are those published essay assignments that are likely to shape teachers' assignments and students' writing.

Throughout the anthologies, each unit test includes essay questions. Across the publishers, there is a range of 1 to 6 essay questions per unit. Most often, the students are given the opportunity to choose one or two questions from among the group of items. If there were not any specific directions, the issue of choice and the number of items to answer would be left up to the teacher.

For the college placement test, students are required to respond to three questions, but this was only one year's examination and the questions change annually. For the state graduation examination, students are asked to respond to one of two literature items. For the state assessment program, students are given one item to which they must respond. However, we were sent about 100 assignments at least some of which are rotated throughout the population, so that it was impossible to determine which would be used at any one time.

Because of the ambiguity of the administration of the essay assignments, it is difficult to assign a precise count of types of essay assignments in a given testing program. What follows is a depiction of the modal assignments in each program.

The categorization system for essay assignments used in this study was adopted from the categories developed by Purves, Söter, Takala and Vähäpassi (1984). Four dimensions of the compositions were analyzed: Stimulus, Cognitive Demand, Purpose, and Rhetorical Specification. Stimulus refers to the nature of the prompt; Cognitive Demand refers to a hierarchical system of intellectual behaviors required to answer the item; Purpose refers to the goal of the writer in answering the question; and Rhetorical Specification refers to the structure or format that the composition should follow. A summary of the categories used appears below and categories with samples of actual assignments appear in Appendix 3.

STIMULUS
- Unspecified
- Word, Phrase, or Sentence
- Literary Text

COGNITIVE DEMAND
- Organize, Reorganize Events
- Organize, Reorganize Ideas
- Invent, Generate, Evaluate

PURPOSE

Unspecified
Multiple
To Learn
To Convey, Signal
To Inform
To Convince, Persuade
To Entertain

RHETORICAL SPECIFICATION

Unspecified
Multiple
Note, Informal Letter
Formal Letter
Summary
Paraphrase
    Narrative
Description
Exposition, Analysis, Definition, Classification
    Narrative, Description, Exposition, with Evaluative Comment
Argument, with Evaluation and Comment
Literary Genre

After two raters practiced using the scheme on a sample of 50 compositions, the raters found that they agreed on all cases. For the main analysis, each composition was scored by a single rater on each dimension and a subsample was check scored by the other rater. There was over 97% agreement in this check scoring.

Results

The results of this analysis are represented in summary form, highlighting the most typified configuration of essay test characteristics.

Textbook Tests

Stimulus:
    Literary Text from prior readings

Cognitive Demand:
    Invent, Generate, Evaluate

Purpose:
    To inform

Rhetorical Specification:

1. Exposition/Analysis
2. Narrative/Exposition with Evaluative Comment

College Placement Test in Literature

Stimulus:

Literary Text from new passages or prior reading

Cognitive Demand:

Invent, Generate, Evaluate

Purpose:

To inform

Rhetorical Specification:

1. Exposition/Analysis
2. Narrative/Exposition with Evaluative Comment

State Graduation Test

Stimulus:

Literary Text from prior reading

Cognitive Demand:

Invent, Generate, Evaluate

Purpose:

To inform

Rhetorical Specification:

Exposition/Analysis

State Assessment Program

Stimulus:

Literary Text from either prior reading or new passages

Cognitive Demand:

Invent, Generate, Evaluate

Purpose:

To Convince or Persuade

Rhetorical Specification:

1. Narrative
2. Argument
3. Note or Informal Letter

Although the anthologies apparently differ from the other essay examinations in never presenting a text with the assignment, this difference simply resulted from the fact that the students were presumed to have the text readily available for reference. The state graduation test and some of the questions on the other tests require that the student draw upon long-term memory in writing about a text that was unavailable. The state assessment differs from the others in consistently asking the student writer to convince or persuade the audience that the writer is knowledgeable or credible. This injunction is often added to other statements of purpose. The same assessment is more varied in the types of text specified and shows a range of formality and text type.

In general, the essay assignments expand upon the objective tests and provide students with a varied selection of topics and cognitive problems with which to deal. The essay assignments tend to be somewhat limited as to the types of writing they engender and they appear to concentrate on the academic essay. There is little encouragement of personal expressive writing or literary writing. Most is informative report writing. It is clear, however, that the essay assignments expand upon the other forms of test and, if used, make the assessment of learning in literature more comprehensive than an examination of multiple-choice tests alone would suggest. A mixture of objective and essay testing would appear to cover a range of cognitive demands and aspects of content, although knowledge of specific texts is largely untested in any form at the secondary level.

Section 4
Summary and Conclusions

This census and content analysis show a remarkable consistency in testing practices with respect to literature. The census suggests that literature, if included in a state assessment, is most frequently included as part of an assessment of reading comprehension. One must conclude from the self-reports that the states, like the proprietary producers of tests, focus their attention on text comprehension at a relatively low level of cognitive processing. They do so without a clear differentiation between reading a literary text and reading a non-literary one.

What this focus means is that a text is viewed as having propositional content that can be easily summarized into a single main idea, point, gist, or theme. This main point is based on the ostensible semantic content of the selection: Thus, a test might include a fictional selection telling of a baseball player who cannot avoid making excuses for himself, even when courting; the story is told through flashbacks by a narrator whose tone is ironic (e.g., Ring Lardner's "Alibi Ike"). Typically, this passage would be treated as an expository text on baseball by the test. There would be occasional items on the literary sorts of features in the selection, but they would generally be token bows to literary demands. The number of items would be insufficient to present a decision-maker with any sort of subscore concerning the student's knowledge of literature, ability to deal with literary texts, or ability to deal with the literary aspects of texts. Yet a variety of such abilities are suggested as important by literary critics and curriculum planners.

Some attention is paid to these abilities in the essay examinations, and literature is perhaps best assessed through an essay examination. In the essay exam the student may use

judgment in selecting what to address and how to address it; also, in essay examinations there may be individual judgments by teachers or raters of the student's ideas, literary acumen, and writing ability. Indeed, such a case has been made by many (e.g., Purves, 1971; Cooper 1986); yet the fact remains that such essay questions seldom occur outside of the literature textbooks, and clearly not in most achievement batteries and state assessments. There the writing is mostly narrative report writing. Writing lucidly about literature and displaying one's ability to discuss literary works critically appears to be limited to the classroom as an option or supplement to other forms of testing and to the examinations for the best students in a limited number of favored states or academic programs.

The tests that now exist in the United States do not live up to the standards set by the examination systems of countries in which student achievement in literature is high. There is little focus on students' abilities to penetrate a text or to use the array of cultural and literary knowledge that has been made available to them. Both of these characterize examinations in other countries; both are called for in the statements of theorists concerned with the literary education of our students (e.g., Scholes, 1985; Hirsch, 1987). The imaginative power of literature and the power of literature to capture the imagination and intellect of the reader remain unexplored in most of these assessments, which treat the texts as if they were no different from articles in encyclopedias. Under these conditions, it would seem difficult for students to see literature as anything but dead and lifeless; this view of literature is perpetuated by the most potent force in the curriculum, the test.

References

Applebee, A., Langer, J., and Mullis, I. (1986). The writing report card. Princeton, NJ: Educational Testing Service.

Cooper, C. (1986). Researching response to literature and the teaching of literature. Ablex.

Cooper, C. (1987). Personal communication concerning the California Writing Assessment.

Cooper, C., and Purves, A. (1973). Responding: A guide to evaluation; Lexington, MA. Ginn.

Grommon, A. L. (ed.) (1976). Reviews of selected published tests in English. Urbana, IL: National Council of Teachers of English.

Hirsch, E. D. (1983). Cultural literacy. The American Scholar. 52, 159-69.

Hirsch, E. D. (1987). Cultural literacy: What every American should know. Boston: Houghton Mifflin.

Langer, J. A., and Applebee, A. N. (1987) How writing shapes thinking. Urbana, IL: National Council of Teachers of English.

Purves, A. C. (1971). Formative and summative evaluation in literature. B. S. Bloom, J. T. Hastings and G. Madaus (Eds.) A handbook of formative and summative evaluation of student learning. New York: McGraw Hill.

Purves, A. C. (1977). Evaluating growth in English. J.R. Squire (Ed.). The teaching of English. The seventy-sixth yearbook of the National Society for the Study of Education. Part I, Chicago, IL: NSSE.

Purves, A. C. (1979a). Evaluation of learning in literature. T. N. Postlethwaite and B. Choppin (Eds.). Evaluation in Education: An International Review Series, 3. Oxford: Pergamon: 1979, 93-172.

Purves, A. C. (1979b). Achievement in reading and literature in the secondary schools: New Zealand in international perspective. Wellington: New Zealand Council for Educational Research and IEA.

Purves, A. C., and Beach, R. C. (1972). Literature and the reader: Research in response to literature, reading interests and the teaching of literature. Urbana, IL: National Council of Teachers of English.

Purves, A. C., Foshay, A. W., and Hansson, G. (1973). Literature education in ten countries. Stockholm: Almqvist and Wiksell.

Purves, A.C., Söter, A., Takala, S., and Vähäpassi, A. (1984). Towards a domain referenced system for classifying composition assignments. Research in the Teaching of English. 18, 385-416.

Scholes, R. (1985). Textual power: Literary theory and the teaching of English. New Haven, CT: Yale University Press.

Appendix 1
Interview Schedule

State Assessment Groups

1. State or Entity:

2. Name and Title of Interviewee:

3. Date of Interview:

4. Is there an assessment or testing program for secondary school students (Grades 7-12)

   Yes, grade or age levels .....

   No, but planned for grade or age levels .....

   No and no plans for one.....

5. What is the frequency of the cycle of testing?

   semi-annual .....

   annual.....

   biennial or longer (how long?).....

6. What aspects of language arts/English are tested?

   Reading .....

   Writing .....

   Language and usage .....

   Literature .....

   Speaking .....

   Listening .....

   Other

   (Specify) .....

7. Is literature learning assessed

as a separate area .....

through a reading assessment .....

through a writing assessment .....

through an assessment of general humanities/social studies ...

8. What aspects of literature learning are assessed?

specific authors and titles .....

literary terms .....

texts or passages .....

content .....

language .....

structure .....

voice .....

folklore and cultural information .....

readers' attributes (attitudes, interests).....

9. In general, what types of item are used? .....

10. What is your score reporting system:

by pupil .....

by class .....

by school .....

by district .....

11. What is the primary audience of the reporting?.....

12. Is your score reporting

norm-referenced? ..... Please describe reference group .....
.....
.....
.....

criterion referenced? ..... Please describe how criterion is established. .....
.....
.....

13. Are your results used for

graduation .....

grade promotion or repetition .....

placement .....

financial purposes (describe) .....
.....
.....

other (describe) .....

14. Could you please send a sample of your tests so that we might do a content analysis? Please send any other descriptive information you might have as well. We will send you a copy of the results of our survey.

Appendix 2

Commercial Achievement Tests Reviewed

Achievement Batteries

1. Basic Skills Assessment Program, 1977
2. California Achievement Test Form E, 1985
3. Comprehensive Assessment Program, 1980
4. Comprehensive Tests of Basic Skills, 1981 & 1983
5. Educational Development Series, 1984
6. Iowa Test of Basic Skills, 1982
7. Metropolitan Achievement Test Advanced 6, Forms L&M, 1984
8. Sequential Tests of Educational Progress (STEP), Levels I&J, 1979
9. SRA Survey of Basic Skills, 1985
10. Stanford Achievement Test Advanced Level, 1981
11. Stanford Test of Academic Skills (TASK), 1982
12. Tests of Achievement and Proficiency, 1986

Reading Tests

1. Gates-McGinitie Reading Tests, 1978
2. Nelson-Denny Reading Test, 1981

High School Level Literature Tests

1. High School Subjects Test in Literature and Vocabulary, 1980
2. Iowa Tests of Educational Development, Form X-7 Literature, 1979

3. Romeo and Juliet computer testpak, Wm. C. Brown Publishers, 1987

College Admissions Tests

1. P-Act, 1987

2. Scholastic Aptitude Test, College Board, 1987

College Placement Tests in Literature

1. Advanced Placement Examination in English Literature and Composition II, 1986 & 1987

2. College Board Achievement Test in Literature, 1981

3. College Level Examination Program

    a. American Literature, 1985

    b. Analysis and Interpretation of Literature, 1987

    c. English Literature, 1987

    d. Humanities, 1986

Literature Anthology Series Publishers

Harcourt Brace Jovanovich

McGraw Hill

Scott Foresman

Scribner Macmillan

Silver Burdett and Ginn

Basal Readers Series Publishers

Harcourt Brace Jovanovich

McGraw Hill

Scott Foresman

Scribner Macmillan

Silver Burdett and Ginn

Appendix 3

Classification Scheme for Test Items

As with any categorization system, not all of the items studied fell clearly within given classifications. When the content of the item included more than one subject area or when it was ambiguously worded, judgments were necessary. It was determined that it was better to try to approximate a classification rather than use a mixed or miscellaneous category.

Sample Items to Illustrate the Classification System

The following are examples to illustrate the type of multiple choice items which were included in some of the categories used to analyze the content of literature items.

Specific author/title at the recall or recognize level:

Which of the following is a play by Anton Chekhov? a) A Slander b) The Wild Duck c) Twelve Angry Men d)The Cherry Orchard.

At the application (Apply) level:

A common theme in much of John Milton's poetry is a) the politics of the times b) the joy of true love c) the centrality of religion d) social inequities.

Literary terms at the recall level:

The turning point of a story is called the a) exposition b) conclusion c) foreshadowing d) climax.

At the apply level:

Lines 24 and 25 are an example of a) onomatopoeia b) alliteration c) personification d) parallelism.

Comprehension (this includes word and section meaning, plot and setting) at the recall level:

Which of the following happens last in the story? a) Tasha walks home b) Bess fantasizes about winning the game c) the girls play softball d) Tasha thinks about summer.

At the describe or summarize level:

What is this passage mostly about? a) daydreaming to escape from problems b) the need to communicate well c) the boy's feelings about his life.

At the apply level:

The phrase "minimal culture" means the people were a) backward b) stagnant c) unsophisticated d) corrupt.

At the analyze level:

The setting of this story was chosen because it a) provides an accurate description b) provides humor for the reader c) it is typical of a horror story d) it provides clues to how the story will end.

Inferences or motives at the recall level:

James went to see Paco because a) his mother asked him to b) he was curious about Paco's new dog c) his friends had dared him to d) he wanted someone to talk to.

At the describe level:

Which of the following best describes the author's main purpose? a) to amuse readers b) to describe an event objectively c) to teach readers a moral d) to persuade the reader

At the apply level:

Which of the following would you most likely have seen Marilyn doing? a) working in the field b) going shopping in the city c) attending the village school d) reading in the loft.

At the interpret level:

The children in the story hope that a) the baby will awaken soon b) their mother will not die c) their father will return d) their mother will bring home enough money.

Characterization at the recall level:

The king was betrayed by a) his wife b) a servant c) the general d) his son.

At the describe level:

Which words best describe George? a) a showoff b) a thoughtful person c) a bully d) a hardworking person.

How did the man feel about the horse? a) excited b) sympathetic c) amused d) terrified.

Theme or moral at the describe level:

The message of this story is a) sorrow and joy are closely related b) humans can not hold time still c) all life has imperfections d) time washes away sorrow.

At the interpret level:

Which sentence best states the poem's theme? a) people are comfortable with routines b) one needs to see new experiences c) conformity d) pursue one goal at a time.

Language (this includes diction, metrics, versification, and figurative language) at the recall level:

The tulips are compared to a) the sun b) people c) fine dresses d) jewels.

At the apply level:

The poet used all of the following devices except a) fixed rhyme scheme b) rapid rhythm c) metaphors d) regular metrical pattern.

At the analyze level:

Which of the following is the best interpretation of "And man flies gasping from what he has made". a) people hide in fear of modern inventions b) people escape from the heat of the city c) people are awed by the sights from the airplane.

Structure (this includes conflict, organization and genre) at the apply level:

The major conflict in the story is between Percy and a) the conformist students b) his fiancee c) his own divergent goals d) his father.

At the analyze level:

What type of literature has this selection been taken from? a) epic poem b) a play c) a short story d) a fable.

Voice at the describe level:

The tone of the passage can best be described as a) peaceful b) hostile c) optimistic d) thankful.

At the apply level:

In contrast to the rest of the poem, the first four lines imply a mood of a) unhappiness b) confusion c) disgust d) anger.

Known topics or cultural information at the recall level:

Prometheus was punished for a) giving humans fire b) making humans stand up like gods c) stealing from the sun d) making animals stronger than man.

At the apply level:

Which best states the similarity between this passage and the fable of the lion and the mouse? a) one character was small and one was large in each b) in both, characters are afraid of each other c) in both, the weaker helps the stronger.

SAMPLES OF ESSAY ITEMS

STIMULUS

Unspecified

Discuss a difficult situation that you have been in and how you overcame it.

Word, Phrase, or Sentence

Discuss the idea: there is nothing to fear but fear itself.

Literary Text

Explain how Fortunato's vanity helped cause his death.

COGNITIVE DEMAND

Organize, Reorganize Events

Select one scene from either, "The Glass Menagerie" or "I Never Sang for My Father", a scene that has a climatic point and write a synopsis of it in short story form. Be sure to include a description of the setting, characters, action.

Organize, Reorganize Ideas

Select two stories from the list and compare and contrast the attitude toward sharing by one character from each story.

Invent, Generate, Evaluate

Explain which character in this unit you respect most and why.

PURPOSE

Unspecified

Compare and contrast Miniver Cheevy and Richard Cory.

To Learn

Setting, the time and place of action is often a key element in short stories. Choose one of the stories you have read in this unit Referring to specific details in the selections, show how the author develops the setting. Then tell why the setting is important to the story.

To Convey, Signal

Discuss any poem that made you feel sad, explaining why.

To Inform

The nineteenth century in America saw the emergence of both Romanticism and realism.

Discuss the most significant characteristics of both literary movements, referring to the work of representative writers.

To Convince, Persuade

Freedom is more important than peace. Agree or disagree.

To Entertain

Write an ending for "August Heat".

Multiple

Pick out two or three lessons that Buck learned; explain them and tell whether or not you believe they really apply in life. (Inform and Convince)

RHETORICAL SPECIFICATION

Unspecified

Explain how Harriet Tubman is valiant

Multiple

Select one of the characters listed and in a paragraph or two, explain the kind of person that character is and the manner in which the playwright reveals his or her qualities. (Description with Evaluative Comment and Analysis)

Note, Informal Letter

Suppose that you are Penelope and that you are writing a letter to Ulysses during his long absence. What will you as Penelope say in your letter?

Formal Letter

Using the ideas expressed in "Declaration of Sentiments," write a letter as Elizabeth Cady Stanton to a public official persuading the official of the justice of your cause.

Summary

The audience learns indirectly about the past of Annie Sullivan. Write a summary in which you tell what you learned about her.

Paraphrase

Discuss any one of the ballad poems, briefly telling the story and explaining the main idea or theme.

Narrative

Tillie Olson's story "I Stand Here Ironing" is an interior monologue written from the mother's point of view. Write an interior monologue from Emily's point of view. Have Emily "talk" about her childhood, her feelings about her family and her achievements.

Description

In about 200 words, describe a place in the city or country, either outdoors or indoors. Write a description that appeals to at least three of the five senses and includes at least one simile. Use enough detail so that the reader will know how you feel about the place you

describe.

Exposition, Analysis, Definition, Classification

Many writers of the 19th century were strongly influenced by the intellectual movement called Transcendentalism. Discuss the major ideas of Transcendentalism and relate these ideas to the worlds you have read by Ralph Waldo Emerson and Henry David Thoreau.

Narrative, Description, Exposition, with Evaluative Comment

Antigone's actions are an example of civil disobedience. She believes that there are times when the laws of man are secondary to higher laws. Write a report justifying or condemning Antigone's act of civil disobedience.

Argument, with Evaluation and Comment

Explain which character in this unit had the greatest handicap and why.

Literary Genre

Write a tall tale like "The Celebrated Jumping Frog of Calaveras County". You may want to take a humorous incident from your own life. Try also to use some of the following humorous techniques: exaggeration, absurdity, caricature, and satire.

Appendix 4
Summary Tables

Appendix Table 1
Commercial Achievement Test Batteries
Reading Comprehension Subtests
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	0.1	0.0	0.0	0.0	0.0	0.0	0.1
Literary Terms	0.0	0.0	0.1	0.0	0.0	0.0	0.1
Comprehension/ Inference	32.0	10.3	24.9	1.5	0.8	0.0	69.3
Characterization	3.1	4.3	2.4	0.0	0.0	0.0	9.8
Theme/Moral	0.0	2.7	1.1	0.1	0.2	0.0	4.1
Language	0.8	0.6	5.9	0.5	0.8	0.0	8.5
Structure	0.2	0.1	2.0	0.6	0.0	0.0	2.8
Voice	0.0	1.0	3.0	0.1	0.0	0.0	4.1
Culture	0.1	0.0	0.4	0.0	0.0	0.0	0.5
Total	36.3	18.8	39.7	2.8	1.8	0.0	100.0

n = 12 test batteries

Appendix Table 2
Literature Content in Reading Tests
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Intepret	Generalize	Total
Author/Title	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Literary Terms	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Comprehension/ Inference	40.0	12.9	15.2	0.0	0.0	0.0	68.1
Characterization	16.1	0.0	3.0	0.0	0.0	0.0	19.1
Language	1.5	0.0	0.0	0.0	0.0	0.0	1.5
Structure	5.7	0.0	0.0	0.0	0.0	0.0	5.7
Voice	0.0	0.0	5.6	0.0	0.0	0.0	5.6
Culture	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	63.2	12.9	23.8	0.0	0.0	0.0	100.0

n = 2 reading tests

Appendix Table 3
High School Literature Tests
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	0.0	0.0	0.7	0.0	0.0	0.0	0.7
Literary Terms	0.0	0.0	1.4	3.4	0.0	0.0	4.7
Comprehension/ Inference	17.8	1.4	29.5	0.7	6.1	0.0	55.4
Characterization	3.0	2.4	5.4	0.7	0.0	0.0	11.5
Theme/Moral	0.0	2.4	0.7	0.0	0.7	0.0	3.7
Language	0.0	0.0	5.8	2.1	1.7	0.0	9.5
Structure	0.0	0.0	8.1	0.7	0.0	0.0	8.7
Voice	0.0	1.4	3.0	1.4	0.0	0.0	5.8
Culture	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	20.9	7.5	54.5	8.9	8.4	0.0	100.0

n = 2 general literature tests

Appendix Table 4
College Admissions Tests
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Literary Terms	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Comprehension/ Inference	31.3	6.3	43.8	0.0	0.0	0.0	81.3
Characterization	0.0	0.0	12.5	0.0	0.0	0.0	12.5
Theme/Moral	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Language	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	0.0	0.0	6.3	0.0	0.0	6.3
Voice	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Culture	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	31.3	6.3	56.3	6.3	0.0	0.0	100.0

n = 2 tests

Appendix Table 5
College Placement Examinations in Literature
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	20.2	1.9	6.2	0.6	0.0	0.3	29.3
Literary Terms	0.2	0.3	1.8	0.2	0.2	0.0	2.7
Comprehension/ Inference	4.4	4.9	12.7	0.4	3.6	0.0	26.1
Characterization	1.7	1.4	2.1	0.3	0.7	0.0	6.2
Theme/Moral	0.7	1.4	1.1	0.2	0.8	0.5	4.7
Language	1.0	8.7	2.1	1.6	0.0	0.0	13.4
Structure	0.2	1.3	3.8	2.4	0.2	0.0	7.9
Voice	0.0	1.4	1.5	0.8	0.0	0.0	3.7
Culture	2.8	0.0	2.9	0.4	0.0	0.0	6.1
Total	31.2	21.4	34.2	7.0	5.5	0.8	100.0

n = 5 placement examinations

Appendix Table 6
High School Textbook Tests, Grade 7
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	7.3	0.2	0.0	1.0	0.0	0.0	8.5
Literary Terms	0.9	0.7	3.7	0.8	0.0	0.0	6.0
Comprehension/ Inference	32.3	1.0	6.1	6.1	0.0	0.0	45.5
Character	18.0	2.0	0.6	1.6	0.0	0.0	22.1
Theme/Moral	7.8	0.8	0.2	0.0	0.0	0.0	8.8
Language	0.3	0.0	0.2	0.4	0.0	0.0	1.0
Structure	2.7	3.5	0.2	2.4	0.0	0.0	8.8
Voice	0.0	0.0	0.0	0.2	0.1	0.0	0.3
Culture	0.1	0.0	0.0	0.0	0.0	0.0	0.1
Total	69.3	8.0	11.1	12.6	0.1	0.0	100.0

Appendix Table 7
High School Textbook Tests, Grade 9
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	3.5	0.1	0.2	0.2	0.0	0.0	3.9
Literary Terms	1.3	1.0	9.8	0.2	0.0	0.0	12.3
Comprehension/ Inference	26.8	0.4	12.2	4.7	0.3	0.0	44.5
Character	10.4	1.5	4.0	2.8	0.0	0.0	18.6
Theme/Moral	4.4	0.5	3.5	0.0	0.0	0.0	8.5
Language	0.7	0.2	2.9	0.2	0.0	0.0	4.1
Structure	3.3	0.6	2.1	0.5	0.0	0.0	6.5
Voice	0.4	0.0	1.0	0.0	0.0	0.0	1.5
Culture	0.9	0.0	0.0	0.0	0.0	0.0	0.9
Total	51.8	4.3	35.6	8.5	0.3	0.0	100.0

Appendix Table 8
High School Textbook Tests, Grade 11
Mean Percent of Items by Content and Cognitive Demand
(Combined Categories)

	Recall	Describe	Apply	Analyze	Interpret	Generalize	Total
Author/Title	5.8	1.2	0.6	4.7	0.0	0.0	12.4
Literary Terms	1.0	0.1	7.0	1.7	0.1	0.0	10.0
Comprehension/ Inference	18.3	3.9	6.7	5.6	0.1	0.0	34.6
Character	6.4	2.6	2.1	2.1	0.0	0.0	13.1
Theme/ Moral	10.5	3.0	3.4	1.1	0.0	0.0	18.0
Language	0.5	0.0	0.5	1.7	0.0	0.0	2.7
Structure	3.1	1.2	0.8	0.9	0.0	0.0	6.0
Voice	0.6	0.4	0.9	0.2	0.2	0.0	2.4
Culture	0.5	0.0	0.2	0.0	0.0	0.0	0.7
Total	46.8	12.4	22.2	18.1	0.4	0.0	100.0